PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 5/10, 9/00, 15/10, A01K 67/027

A2

(11) International Publication Number: **WO 99/53031**

(43) International Publication Date: 21 October 1999 (21.10.99)

(21) International Application Number: PCT/EP99/02391

(22) International Filing Date: 8 April 1999 (08.04.99)

(30) Priority Data:
198 15 590.5 8 April 1998 (08.04.98) DE

(71) Applicant *(for all designated States except US)*: MAX-PLANCK-GESELLSCHAFT ZUR FÖRDERUNG DER WISSENSCHAFTEN E.V. [DE/DE]; Berlin (DE).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: KURZCHALIA, Teymuras [GE/DE]; Fehrbelliner Strasse 97, D-10119 Berlin (DE).

(74) Agent: VOSSIUS & PARTNER; Siebertstrasse 4, D-81675 München (DE).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, GI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: PROCESS FOR PRODUCING CELL CLONE LIBRARIES

(57) Abstract

Described is a process for producing cell clone libraries via the use of a pool of expression vectors comprising ribozyme genes.Furthermore, cell clones with modified phenotypes and non-human animals derived from such cell clones are provided as well as novel genes identified from the cell clone library. Fields of the application of the invention are molecular-biological research and the pharmaceutical industry.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

Process for producing cell clone libraries

The invention relates to a process for producing cell clone libraries. Fields of application of the invention are molecular-biological research and the pharmaceutical industry.

Cell clones having properties that are modified as compared to natural cell populations play a substantial role in molecular-biological research. For example, in particular cells lacking certain genes (deficient cells) are widely used to recognize molecular mechanisms. A classical example thereof is the identification of 33 genes which are required for posttranslational events in the yeast secretion pathway (P. Nowick et al., Cell, Vol. 21 (1980), 205-215). For somatic cells, it is more difficult to use genetic methods (M. Krieger, Methods in Enzymology 129 (1986), 227-237). Recently, modifications to embryonic stem cells (ES cells) have been made, making available for tests test animals grown therefrom which exhibit a known gene deficit (knock-out mice). However, it requires a lot of time and/or money to develop such knock-out mice and requires a minimum time of 1 to 1.5 years.

Thus, there is a need, particularly in the pharmaceutical industry, for having at one's disposal not only single cell clones but a large range of clones to choose from.

The problem underlying the invention is therefore to make cell clone libraries available. It is based on the object of developing a simple process for developing such libraries.

The problem underlying the invention is solved by the embodiments characterized in the claims.

Accordingly, the present invention relates to a method for producing cell clone libraries comprising the steps of:

(a) transfecting mammalian cells with a pool of expression vectors comprising

(i) ribozyme genes which comprise a central catalytic domain and flanking sequences of at least 6 nucleotides in length each of random base sequence and, optionally,

(ii) a selection marker;

(b) cultivating the transfected cells under conditions which permit expression of the ribozyme gene present in the vectors;

(c) identifying those cell clones expressing the ribozyme;

(d) optionally subdividing the vector pool(s) identified in step (c) and repeating step (a) to (c); and

(e) examining the so-identified cell clones for a modification of the properties they originally had.

"Ribozymes" are RNA molecules possessing specific catalytic activities, i.e. endoribonuclease activities. The most investigated ribozyme sequences are based on the catalytic domain derived from a satellite RNA sometimes associated with tobacco ring spot virus. Due to a potential folding conformation, such sequences are termed "hammerhead" ribozymes (Forster and Simons, Cell 49 (1987), 211-220). The catalytic sequence derived from these hammerhead ribozymes can be modified and directed to cleave a new RNA by designing flanking sequences that are complementary to the new target RNA (Haseloff and Gerlach, Nature 334 (1988), 585-591). Provided adequate complementarity between the target and ribozyme RNA is present, the minimum requirement for efficient cleavage is a 3-base sequence. To date, in vitro analysis has determined that the triplets GUC, GUU, GUA, CUC, or UUC on the RNA provide efficient target sites, with cleavage occurring following the triplet. Other triplets, particularly those with a uracil as the central residue, may also prove to be effective target sites. Ribozymes can therefore be designed to target any given RNA sequence provided that it contains the necessary target site triplet.

In the examples of the present invention the target triplet GUC was used although as mentioned before other target triplets can be employed as well. Furthermore, the present invention is not restricted to the use of the hammerhead ribozymes but can also be performed with the so-called "hairpin" ribozymes described in EP-A1 0 321 201 and with other ribozymes that have meanwhile been developed.

The terms "catalytic domain" or "catalytic region" are used interchangeably herein and denote a base sequence which is adapted to cleave a given target RNA. Such catalytic regions and corresponding ribozymes and ribozyme genes are described in the prior art, for example in EP-A1 0 321 201 and EP-A1 0 360 257, the disclosure content of which is incorporated herein by reference. Further catalytic domains and catalytic regions can be employed from other natural and synthetic RNA catalysts described in the prior art or that will be developed, for example in accordance with the basic teaching in WO88/04300, the disclosure content of which is hereby incorporated by reference.

The "flanking sequences" define a hybridizing region comprising one or more arms formed of single-stranded RNA and having a sequence complementary to at least part of a target RNA, said one or more arms being associated with the catalytic domain capable of cleaving said target RNA; and where the flanking sequences provide a single arm of RNA, said arm contains at least 9 nucleotides and where the flanking sequences comprises two arms of RNA, the sum of nucleotides in each of said arm is greater than 6 nucleotides.

While in the prior art ribozymes have been used to specifically inactivate a desired predetermined target RNA molecule, in the method of the present invention a vector pool of different ribozyme genes is generated by randomly selecting the nucleotides of the flanking sequences; see also the appended examples.

Step (e) of examining identified cell clones for a modification of the properties they originally had refers to the determination of a responsive change in the cell clones containing the ribozyme genes compared to corresponding wildtype cell cultured in the absence of a ribozyme gene. In this context the term "wildtype cell" refers to a mammalian cell prior to transfection with the ribozyme genes. Said mammalian cells may either be naturally occurring mammalian cells or genetically modified, i.e. transgenic or transformed mammalian cells or can be mutant mammalian cells obtainable by the exposure to chemical or physical treatment such as application of nitrosocompounds and radiation, respectively. Various methods for obtaining such modified mammalian cells are well known to the person skilled in the art and are described, e.g., in Sambrook et al. or Ausubel et al.; see infra.

The term "modification of the properties [the identified cell clones] originally had" or "responsive change" refers to a change in a phenotypic characteristic of said cells

transfected with the ribozyme genes other than the expression of the ribozyme genes in said cells per se and includes graded cellular responses, preferably observable with the eye such as cultural or morphological responses as well as changes in signal transduction pathways such as receptor activity which can be measured by, e.g., in vitro binding and/or activity assays. Preferably, the response that has changed is apoptotic cell death or morphology of the cell or polarization of epithelial cells. The principles of determining the modification of the properties of cells are well known to the person skilled in the art and are described, for example, in EP-A1 0 403 506 and Current Protocols in Cell Biology, John Wiley and Sons, New York (1997). As described in the appended examples, the transfected cells can be transplanted or injected into animals such as mice, preferably after culturing in a selective medium for 10 to 14 days. As control animals injected with the corresponding wildtype cells are monitored and compared to the animals comprising the transfected cell clones. Monitoring can be done, for example, by determining reporter gene expression or by monitoring behavior of the "transgenic" non-human animals and the wildtype animals. Animals employed in accordance with the present invention include but are not limited to mouse, fish, Caenorhabditis elegans, Drosophila melanogaster.

The conditions of transfecting the mammalian cells with the pool of expression vectors may depend on various factors, e.g., excess/deficiency, the duration of incubation, sorting method, etc. and can be adjusted using standard protocols known in the art for transfection of mammalian cells. Such transfection protocols are described in, for example, Current Protocols in Molecular Biology, John Wiley and Sons, New York (1997).

The principle of expression cassettes suitable for ribozyme gene expression is described, for example, in DE-C1 44 27 7061 the disclosure content of which is hereby incorporated by reference.

In accordance with the present invention, it was surprisingly found that cell clone libraries can be developed by using a library or vector pool containing ribozymes. In the examples of the present invention, these ribozyme genes consist of a central hammerhead structure from double-stranded DNA of the sequence

5'-CTGATGAGTCCGTGAGGACGAAAAC-3' (SEQ ID NO: 1) with flanking sequences to both sides of the hammer-head structure of 6 to 13 nucleotides each of random base sequence. They furthermore contain an antibiotics selection marker. According to the invention, this ribozyme library was used for the transfection of cells with certain properties, whereupon ribozyme-expressing cell clones are permanently selected by treatment with the antibiotic and are examined for a modification of the properties they originally had. In this transfection, a part of the cells used is killed and the ribozyme genes in particular their flanking sequences are identified that are responsible for this process.

As demonstrated in Example 1, several cDNAs encoding known and putative tumor suppressor genes could be identified and cloned from MDCK cells. Similarly, as is described in Example 3, this approach can be easily used for identification of intracellular lipid/protein transport proteins. In fact, the present invention provides a generally applicable method for the identification of genes and proteins which give rise to a phenotype at least under certain conditions, for example if there are over-/underexpressed either in cell culture or in an animal model system as described in Example 1. Thus, in a preferred embodiment of the method of the present invention, the cultivation in step (b) is performed in an animal, preferably mice; see, e.g., Gene Targeting, Ed. A.L. Joyner, IRL Press, Oxford University Press, Oxford, New York Tokyo (1993).

As described in Example 1, the method of the present invention is particularly suited for the identification of tumor suppressor genes. For this purpose, it is recommendable to examine the cell clone library established by the method of the present invention for the modification for important biological activities of a tumor suppressor. Examples for important biological activities of a tumor suppressor are the capability to inhibit in-vitro proliferation of tumor cells as evidenced for instance by measuring colony formation, growth rate and cloning in soft agar as well as the capability to inhibit in-vivo tumor formation in nude mice. These biological activities can be determined, for example, according to Zhou et al., Proc. Natl. Acad. Sci. USA 91 (1994), 4165-4169; Chen et al., Science 250 (1990), 1576-1580; Baker et al., Science 249 (1990), 912-915; Diller et al., Mol. Cell. Biol. 10 (1990), 5772-5781; Casey et al., Oncogene 6 (1991), 1791-1797; Cheng et al., Cancer Research (1992),

222-226; Wang and Prives, Nature 376 (1995), 88-91; Mercer et al., Proc. Natl. Acad. Sci. USA 87 (1990), 6166-6170; Antelman et al., Oncogene 10 (1995), 697-704 or as described in the appended examples.

The term "pool of expression vectors" or "vector pool(s)" in step (b) of the method according to the invention is meant to be understood as a plurality of vector molecules which are preferably not identical and which are adapted for expression in the transfected cells. Such vector molecules comprise regulatory elements which are capable of directing expression of a linked sequence, i.e. the ribozyme genes in the transfected cells. The vectors of said vector pool may furthermore comprise sequences which ensure replication in prokaryotic host cells as well as sequences which ensure replication in the transfected eukaryotic cells. The vectors employed in accordance with the method of the invention may be designed for direct introduction or for introduction via liposomes, or viral vectors (e.g. adenoviral, retroviral) containing said recombinant DNA molecule into the cell. Preferably, said cell is a germ line cell, embryonic cell, stem cell or egg cell or derived therefrom. An embryonic cell can be for example an embryonic stem cell as described in, e.g. Nagy, Proc. Natl. Acad. Sci. USA 90 (1993), 8424-8428.

It is to be understood that the introduced vectors express either upon induction or constitutively the ribozyme gene after introduction into said cell and preferably remain in this status during the lifetime of said cell. For example, cell lines which stably express the ribozyme under the control of appropriate regulatory sequences may be engineered according to methods well known to those skilled in the art. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with the vector of the invention and a selectable marker, either on the same or separate vectors. Suitable regulatory elements, i.e. promoters that can be used to drive the expression of the ribozyme gene include viral promoters like RSV, MPSV, SV40, CMV or 7.5k, vaccinia promoter, inducible promoters, metalothionein promoter, etc.

In this context, suitable expression vectors are known in the art such as Okayama-Berg cDNA expression vector pcDV1 (Pharmacia), pCDM8, pRc/CMV, pcDNA1, pcDNA3 (In-vitrogene), pSPORT1(GIBCO BRL).

If a vector pool is identified in step (c) of the method according to the invention then it is either possible to isolate from the original pool of the so-identified vector pool the

ribozyme sequences present in the vectors of the vector pool and characterize the flanking sequences or one can further subdivide the original vector pool, for example, if it consists of vectors with a plurality of different ribozyme genes, so as to reduce the number of different vectors per pool and repeat the method with the subdivisions of the original pool. Depending on the complexity of the pool this can be done for several times, preferably so often until the vector pool identified in step (c) of the method only comprises a limited number of vectors which differ with respect to their inserted ribozyme gene. Normally the vector pool used in step (a) for transfecting the cells has previously been isolated from a mixture of bacteria harboring different vectors and which, thus, constitute a kind of library. Subdivision of the vector pool for the purpose of step (c) of the method can therefore be achieved by subdividing said library comprising the bacteria so that the diversity of the vectors with respect to the inserted nucleic acid sequences is lower in the subdivisions than in the original library. From these subdivisions of the library the expression vectors can then be isolated. These isolates then represent subdivisions of the original vector pool.

The ribozyme genes present in vectors of a vector pool identified by (c) of the method according to the invention can be isolated from the vectors, e.g. by digestion with suitable restriction enzymes and can be further characterized, for example by restriction mapping, sequencing etc.

In a preferred embodiment of the invention the mammalian cells are Madin-Darby Canine Kidney (MDCK, ATCC-collection #CCL-34), mouse embryonic stem cells (ES, ATCC collection, #CRL-2239), CHO-cells (ATCC-collection, #CCL-61) or other embryonic stem cells. Other suitable mammalian cell lines comprise Caco-2 (human colon carcinoma, ATCC #CRL-2102 and #HTB-37), 293 (human embryonal kidney, ATCC # CRL-1573), Hela (human adenocarcinoma, ATCC # CCL-2), 3T3 (mouse embryonic firbroblast, ATCC # CCL-163), primary endothelial cell cultures, Saos-2 human osteosarcoma cells (ATCC HTB-85), HeLa human epidermoid carcinoma cells (ATCC CRL-7923), HepG2 human hepatoma cells (ATCC HB-8065), human fibroblasts (ATCC CRL-1634), U937 human histiocytic lymphoma cells (ATCC CRL-7939), RD human embryonal rhabdomyosarcoma cells (ATCC CCL-136), MCF7 human breast adenocarcinoma cells (ATCC HTB-22), JEG-3 human

choriocarcinoma cells (ATCC HB36), A7r5 fetal rat aortic smooth muscle cells (ATCC CRL-1444), and NIH 3T3 mouse fibroblasts (ATCC CRL-1658) obtainable from the American Type Culture Collection. The method of the present invention can be generally conducted on primary cultures including, but not limited to, a smooth muscle cell, an endothelial cell, a basophil, a mast cell, an eosinophil, a neutrophil, a macrophage, a B lymphocyte, a T lymphocyte, a dendritic cell, a natural killer cell, a plasma cell, a neuroblast, a stem cell or a cancer cell.

In one embodiment of the method of the present invention the regulatory elements controlling the expression of the ribozyme genes comprise a constitutive, tissue-specific, organ-specific or inducible promoter including estrogen activated promoter (Reichmann et al., Cell 24 (1992), 1103); promoter of mouse Mx1 gene inducible by interferon (Kühn et al., Nature 269 (1995), 14279), heat shock promoter for studies in Drosophila or C. elegans (described for example in: Nematode Caenorabditis elegans, Cold Spring Harbor Laboratory Press, Ed. W. Wood 1988). A suitable inducible system is for example tetracycline-regulated gene expression which is described by, e.g., Gossen (Proc. Natl. Acad. Sci. USA 89 (1992), 5547-5551; Trends Biotech. 12 (1994), 58-62). Further inducible expression systems comprise Ectyson-system (Invitrogen), Lac-repressor system (Stratagen) and Tet-system (Clontech).

As mentioned above, in the method of the present invention the vectors may contain a selection marker, e.g., an antibiotics selection marker or a reporter gene, which are advantageously used since they allow rapid screening and/or selection of transfected cells.

For example, following the introduction of foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines.

A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler, Cell 11(1977), 223), hypoxanthine-guanine phosphoribosyltransferase (Szybalska, Proc. Natl. Acad. Sci. USA 48 (1962), 2026), and adenine phosphoribosyltransferase (Lowy, Cell 22 (1980), 817) genes can be employed in tk^-, $hgprt^-$ or $aprt^-$ cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for dhfr, which confers resistance to methotrexate (Wigler, Proc. Natl. Acad. Sci. USA 77 (1980), 3567; O'Hare, Proc. Natl. Acad. Sci. USA 78 (1981), 1527), gpt, which confers resistance to mycophenolic acid (Mulligan, Proc. Natl. Acad. Sci. USA 78 (1981), 2072); neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, J. Mol. Biol. 150 (1981), 1); and hygro, which confers resistance to hygromycin (Santerre, Gene 30 (1984), 147) genes. Additional selectable genes have been described, namely trpB, which allows cells to utilize indole in place of tryptophan; hisD, which allows cells to utilize histinol in place of histidine (Hartman, Proc. Natl. Acad. Sci. USA 85 (1988), 8047); and ODC (ornithine decarboxylase) which confers resistance to the ornithine decarboxylase inhibitor, 2-(difluoromethyl)-DL-ornithine, DFMO (McConlogue, 1987, In: Current Communications in Molecular Biology, Cold Spring Harbor Laboratory ed.).

The reporter gene that can be present in the vector may be any suitable reporter gene the expression of which can be detected in the transfected cells. Preferably, a reporter gene is chosen the expression of which can be easily detected, for example, by photometric or fluorometric methods, isotopic labeling or by a staining reaction. Examples for reporter genes preferably used in the method according to the invention are those coding for chloramphenicol-acetyltransferase (CAT), ß-galactosidase (ß-Gal), secreted alkaline phosphatase (SEAP) or growth hormone (GH).

Advantageously, said marker is a gene encoding luciferase (Giacomin, Pl. Sci. 116 (1996), 59-72; Scikantha, J. Bact. 178 (1996), 121), green fluorescent protein (Gerdes, FEBS Lett. 389 (1996), 44-47) or ß-glucuronidase (Jefferson, EMBO J. 6 (1987), 3901-3907). This embodiment is particularly useful for simple and rapid screening of cells, tissues and organisms containing a vector of the invention.

In a preferred embodiment of the method of the present invention, the antibiotics selection marker encodes neomycin phosphotransferase. The neomycin phosphotransferase gene confers antibiotic resistance in mammalian cells which can be easily selected for the stable dominant marker. For example, many mammalian cell lines, such as mouse 3T3, LCV1, TC7, and hela cells, are sensitive to 100 µg/ml of G418. The use of 400 to 800 µg/ml is recommended; see also the appended examples.

In the method of the present invention the selected properties can be detected by fluorescence analysis, wherein the cells are optionally enriched by cell sorting. This embodiment is particularly advantageous for the rapid selection of cell clones; see example 3. The automatical procedure could be repeated many times leading to further enrichment of desired clones. This method also can be used for studying cells expressing Green Fluorescent Protein (GFP) as a marker for cellular localization, transport or morphology.

In a preferred embodiment of the method of the present invention, the flanking sequences in the ribozyme gene are 6 to 13 nucleotides in length. Generally, the flanking sequences of the ribozyme gene should be sufficient in length in order to allow the ribozyme to be targeted to a unique nucleic acid molecule in the cell. Considering the average complexity of the mammalian genome of about 3×10^9 only about 10% of which represents virtually important sequences while the major portion of the mammalian genome consists of repetitive DNA elements and satellite DNAs flanking sequences a sum of 15 nucleotides in length would statistically sufficient to specifically target one nucleic acid molecule, i.e., mRNA in the mammalian genome. Thus, in case flanking sequences comprise two arms of RNA, both arms may be advantageously 7 or 8 nucleotides in length, i.e. 14, 15 or 16 nucleotides together. Each combination of the length of the arms of flanking sequences is possible, thus ranging from 11 to 26 nucleotides. Of course, it is possible to also use longer flanking sequences, e.g., 6 to 15, 20, 25 or even more than 30 nucleotides in length although less preferred.

As demonstrated in the appended examples, the catalytic domain in the ribozyme gene preferably consists of a hammerhead structure, for example, wherein said hammerhead structure has the sequence 5'-CTGATGAGTCCGTGAGGACGAAAAC-3' (SEQ ID NO: 1)

Furthermore, as described in the appended examples the phenotype-determining ribozyme in the permanently expressing cell clones is amplified by RT-PCR, cloned into a vector and identified by sequence analysis. Then, the respective protein is concluded from the phenotype-determining ribozyme either by data bank analysis or by cloning from cDNA libraries.

In principle, the phenotype-determining ribozyme, i.e. the ribozyme that is responsible for the modification of the properties of the mammalian cell is amplified by amplification techniques such as polymerase chain reaction (PCR). The ribozyme sequences, in particular the flanking sequences of the ribozyme can either be amplified from the ribozyme gene present in selected cell clones or from RNA extracted from the selected cell clone, wherein the RNA is converted into cDNA by reverse transcription and then amplified by, e.g, PCR. Such amplification techniques are well known to the person skilled in the art and are described in, e.g., PCR, Gassen et al., Gustav Fisher Verlag Stuttgart, 1994.

The nucleotide sequence of the flanking sequences of the ribozyme can then be identified by sequence analysis; see Sambrook et al. or Ausubel et al., infra. Although not necessary, the amplified ribozyme sequences may be conveniently cloned in a vector such as pUC or other sequencing vectors known in the art and available from various sources.

However, also methods exist that allow direct sequencing of DNA or RNA extracted from the cells; see Current Protocols in Molecular Biology, John Wiley and Sons, New York, 1997.

Having determined the nucleotide sequence of the phenotype-determining ribozyme the target nucleic acid molecule of said ribozyme can be deduced as said flanking sequences are complementary or at least substantially complementary to the nucleotide sequence of the target nucleic acid molecule. The deduced nucleotide sequence of the target nucleic acid molecule can then be used, for example, to identify the corresponding gene by databank analysis and/or for the generation of

specific probes for the cloning of a gene corresponding to the target nucleic acid molecule from DNA libraries by hybridization. On the other hand, it may not even be necessary to sequence the flanking sequences of the phenotype-determining ribozyme but it is also possible to use said flanking sequences either isolated, combined together or with the whole ribozyme gene as a hybridizing probe. Hybridization and labeling techniques are well known to the person skilled in the art and are described, for example, in Sambrook et al. or Ausubel et al., see infra. For example, nucleic acid probes may be labeled with commonly employed radioactive labels such as ^{32}P and ^{35}S although other labels such as biotin or mercury or chemoluminescent markers may be employed as well. Various methods were known to a person skilled in the art may be used to label a nucleic acid probe. For example DNA sequences and RNA sequences may be labeled with the mentioned labels using the random primer method. Various methods for detection of nucleic acid molecules are also well known in the art, e.g., Southern and Northern blotting, PCR, primer extension and the like.

By "hybridizing" it is meant that such nucleic acid molecules hybridize under conventional hybridization conditions, preferably under stringent conditions such as described by, e.g., Sambrook (Molecular Cloning; A Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989)). An example of one such stringent hybridization condition is hybridization at 4XSSC at 65 °C, followed by a washing in 0.1XSSC at 65 °C for one hour. Alternatively, an exemplary stringent hybridization condition is in 50 % formamide, 4XSSC at 42 °C. Examples of such non-stringent hybridization conditions are 4XSSC at 50 °C or hybridization with 30-40 % formamide at 42 °C.

In a preferred embodiment of the method of the present invention said DNA library is a cDNA library. In particular, said cDNA may be prepared from RNA obtained from any organism or tissue, namely from any animal, bacterial, fungal or plant cells or from viruses. Most preferably, the RNA is obtained form mammalian cells. In such case the RNA is preferably derived from a specific tissue or organ of a mammal, e.g., intestine, stomach, lung, adrenal gland, kidney, mammary gland, prostate, testis, most preferably said tissue is pituitary gland, brain or ovary.

Extending the polynucleotide sequence obtained in accordance with the invention

The polynucleotide sequences identified in accordance with the method of the present invention can be extended utilizing partial nucleotide sequence and various methods known in the art to detect upstream sequences such as promoters and regulatory elements. Gobinda, (PCR Methods Applic. 2 (1993), 318-322) discloses "restriction-site" polymerase chain reaction (PCR) as a direct method which uses universal primers to retrieve unknown sequence adjacent to a known locus. First, genomic DNA is amplified in the presence of primer to a linker sequence and a primer specific to the known region. The amplified sequences are subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR can be used to amplify or extend sequences using divergent primers based on a known region (Triglia, Nucleic Acids Res. 16 (1988), 8186). The primers may be designed using OLIGO® 4.06 Primer Analysis Software (1992; National Biosciences Inc, Plymouth MN), or another appropriate program to be preferably 22-30 nucleotides in length, to have a GC content of preferably 50% or more, and to anneal to the target sequence at temperatures preferably about 68°-72°C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

Capture PCR (Lagerstrom, PCR Methods Applic. 1 (1991), 111-119) is a method for PCR amplification of DNA fragments adjacent to a known sequence in, e.g., human or plant yeast artificial chromosome DNA. Capture PCR also requires multiple restriction enzyme digestions and ligations to place an engineered double-stranded sequence into an unknown portion of the DNA molecule before PCR.

Another method which may be used to retrieve unknown sequences is that of Parker, (Nucleic Acids Res. 19 (1991), 3055-3060). Additionally, one can use PCR, nested primers and PromoterFinder libraries to walk in genomic DNA (PromoterFinder™ Clontech (Palo Alto CA). This process avoids the need to screen libraries and is useful in finding intron/exon junctions. Preferred libraries for screening for full length cDNAs are ones that have been size-selected to include larger cDNAs. Also, random

primed libraries are preferred in that they will contain more sequences which contain the 5' and upstream regions of genes. A randomly primed library may be particularly useful if an oligo d(T) library does not yield a full-length cDNA. Genomic libraries are useful for extension into the 5' nontranslated regulatory region. Capillary electrophoresis may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products; see, e.g., Sambrook, infra. Systems for rapid sequencing are available from Perkin Elmer, Beckmann Instruments (Fullerton CA), and other companies.

Computer-assisted identification of proteins and their encoding genes

Computer programs such as BLAST2, which stands for Basic Local Alignment Search Tool (Altschul, 1997; Altschul, J. Mol. Evol. 36 (1993), 290-300; Altschul, J. Mol. Biol. 215 (1990), 403-410), can be used to search for local sequence alignments. BLAST produces alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST is especially useful in determining exact matches or in identifying homologs. The fundamental unit of BLAST algorithm output is the High-scoring Segment Pair (HSP). An HSP consists of two sequence fragments of arbitrary but equal lengths whose alignment is locally maximal and for which the alignment score meets or exceeds a threshold or cutoff score set by the user. The BLAST approach is to look for HSPs between a query sequence and a database sequence, to evaluate the statistical significance of any matches found, and to report only those matches which satisfy the user-selected threshold of significance. The parameter E establishes the statistically significant threshold for reporting database sequence matches. E is interpreted as the upper bound of the expected frequency of chance occurrence of an HSP (or set of HSPs) within the context of the entire database search. Any database sequence whose match satisfies E is reported in the program output.

Analogous computer techniques using BLAST (Altschul, 1997, 1993 and 1990, supra) are used to search for identical or related molecules in nucleotide databases such as GenBank or EMBL. This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be

modified to determine whether any particular match is categorized as exact or homologous. The basis of the search is the product score which is defined as:

$$\frac{\%\text{sequence identity} \times \%\text{ maximum BLAST score}}{100}$$

and it takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1-2% error; and at 70, the match will be exact. Homologous molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

In a preferred embodiment of the method of the present invention said target nucleic acid molecule encodes a protein. The amino acid sequence of said protein can be analyzed in the databank. For example, a homology search can be performed with the program BLASTX and BLASTN (version 2.0a19MP-WashU [build decunix3.2 01:53:29 05-feb-1998] (see Altschul, Nucleic Acids Res. 25 (1997), 3389-3402) on the appropriate nucleic acids and protein database at National Center for Biotechnology Information, Washington, USA, (http://www.ncbi.nlm.nih.gov); The Institute for Genomic Research (TIGR); The Sanger Centre, Hinxton, Cambridge, UK (http://www.sanger.ac.uk). The function GAP (general alignment) (from the GCG 9.1 package, Genetics Computer Group Inc., Madison, USA) can be used with the parameters Gap weight = 12 and Length weight = 4 to quantify the percentage of homology and similarity. The protein sequences can then be used to perform a BLASTP (version 2.0.4 [feb-24-1998]) with BEAUTY post-processing provided by the Human Genome Center, Baylor College of Medicine against the National Center for Biotechnology Information's non-redundant protein database (http://dot.imgen.bcm.tmc.edu:9331/seq-search/protein-search.html).

The expression products of the identified nucleic acid molecules can be furthermore characterized by expressing them in prokaryotic or eukaryotic host cells and purifying them. Subsequently, enzymatic and/or other biological activities can be determined by in-vitro assays. Expression in eukaryotic host cells or in-vitro transcription and translation systems may furthermore provide information about possible phoshporylation and/or glycosylation patterns etc.

Furthermore, the present invention relates to cell clone libraries obtainable by the method of the present invention described above. Preferably, said cell clone library consists of embryonic stem cells, preferably of mice. Of particular importance is the application of the process according to the invention for embryonic stem cells (ES cells), particularly ES cells of mice. Clone libraries of this kind may be used to specifically breed deficient mouse strains, which considerably rationalizes the search for new pharmaceuticals. Thus, the stem cells obtained in accordance with the present invention can be advantageously used to breed test animals, preferably mice.

The present invention also relates to a method for the production of a transgenic animal, preferably transgenic mice, comprising the method of the present invention, wherein the vector pool is introduced into a germ cell, an embryonic cell, stem cell or an egg or a cell derived therefrom. The non-human animal to be used in the method of the invention may be a non-transgenic healthy animal, or may have a disease, e.g. cancer. Production of transgenic embryos and screening of those can be performed, e.g., as described by A. L. Joyner Ed., Gene Targeting, A Practical Approach (1993), Oxford University Press. The DNA of the embryonal membranes of embryos can be analyzed using Southern blots with an appropriate probe. The invention also relates to transgenic non-human animals such as transgenic mice, rats, hamsters, dogs, monkeys, rabbits or pigs comprising embryonic stem cells as described above or obtained by the method described above, preferably wherein said vector is stably integrated into the genome of said non-human animal, preferably such that the presence of said vector leads to the expression of the ribozyme gene(s). The generation of knockout mice is described in Thomas and Capecchi (Cell 51, (1987), 503-512).

The present invention also relates to nucleic acid molecules obtainable by the method of the invention defined above, wherein said nucleic acid molecule encodes a protein which preferably corresponds to a gene selected from the group consisting of tumor suppressors, genes involved in apoptosis, genes involved in the organization of cytoskeleton, activators and inhibitors of cell migration, genes

involved in metastasis, genes involved in protein trafficking in the cell, genes involved in the polarization of epithelial cell, and genes responsible for resistance of mammalian cells to pathogenic infection. These genes represent the preferred targets the method is employed for. It is evident to the person skilled in the art that once such a nucleic acid molecule has been identified and isolated it is possible to modify the encoded protein by for example, amino acid deletions, substitutions, and/or additions in order to arrive at the mutated version of said protein which has lost its biological activity. Therefore, the present invention also relates to nucleic acid molecules which specifically hybridize with the nucleic acid molecule identified in accordance with the present invention and which encode mutated version of the protein encoded by the originally identified nucleic acid molecule which has lost its biological activity. Hereby, the activity could be an enzymatic (protease, nuclease, involved in metabolic pathways etc); loss of interaction ability with other proteins leading to impairment of cell-cell interactions, disruption of the cytoskeleton or protein trafficking; loss of regulatory activity e.g. aberrant cell cycle or transcriptional activation.

As demonstrated in the examples of the present invention, the above-described method is particularly suitable to identify and isolate nucleic acid molecules which encode proteins having the biological activity of a tumor suppressor or an intracellular lipid/protein transport protein.

Thus, in another aspect the present invention relates to nucleic acid molecules obtainable by a method according to the invention which encode a protein or polypeptide having tumor suppressor activity or intracellular lipid/protein transport protein activity. In a preferred embodiment the nucleic acid molecules according to the invention are DNA molecules, most preferably cDNA molecules.

In a preferred embodiment the nucleic acid molecules according to the invention are derived from a mammal, most preferably form a human or a mouse.

With the help of nucleic acid molecules identified and isolated by the method according to the invention it is possible to isolate the same or related molecules from the same or different organisms, for example, by screening genomic or cDNA libraries with the nucleic acid molecules isolated according to described method as a probe.

Detection and mapping of related polynucleotide sequences

The nucleic acid sequences identified in accordance with the method of the invention can also be used to generate hybridization probes for mapping the naturally occurring genomic sequence. The sequence may be mapped to a particular chromosome or to a specific region of the chromosome using well known techniques. These include *in situ* hybridization to chromosomal spreads, flow-sorted chromosomal preparations, or artificial chromosome constructions such as yeast artificial chromosomes, bacterial artificial chromosomes, bacterial P1 constructions or single chromosome cDNA libraries as reviewed in Price (Blood Rev. 7 (1993), 127-134) and Trask (Trends Genet. 7 (1991), 149-154). The technique of fluorescent *in situ* hybridization of chromosome spreads has been described, among other places, in Verma, (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY. Fluorescent *in situ* hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in The Sanger Centre, Hinxton, Cambridge, UK (http://www.sanger.ac.uk). Correlation between the location of the gene that has been affected by the expression of the phenotypic-determining ribozyme in the method of the invention on a physical chromosomal map and a specific feature, e.g., cell growth, disease etc. may help delimit the region of DNA associated with this feature. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals. Furthermore, the means and methods described herein can be used for marker-assisted breeding of animals.

In situ hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers may be used for extending genetic maps. For example an sequence tagged site based map of the human genome was recently published by the Whitehead-MIT Center for Genomic Research (Hudson, Science 270 (1995), 1945-1954). Often the placement of a gene on the chromosome of another species may reveal associated marker even if the number or arm of a particular chromosome is not known. New sequences can be

assigned to chromosomal arms, or parts thereof, by physical mapping. This provides valuable information to investigators searching for interacting genes using positional cloning or other gene discovery techniques. Once such gene has been crudely localized by genetic linkage to a particular genomic region, any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier or affected individuals.

Vectors and expression systems

The present invention also relates to vectors, particularly plasmids, cosmids, viruses, bacteriophages and other vectors used conventionally in genetic engineering that contain a nucleic acid molecule obtained according to the method of the invention. Methods which are well known to those skilled in the art can be used to construct various plasmids and vectors; see, for example, the techniques described in Sambrook, Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory (1989) N.Y. and Ausubel, Current Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y. (1989), (1994). Plasmids and vectors to be preferably employed in accordance with the present invention include those well known in the art. Alternatively, the nucleic acid molecules and vectors of the invention can be reconstituted into liposomes for delivery to target cells.

Furthermore, the invention relates to a host cell comprising the above-described vector. The host cell can be any prokaryotic or eukaryotic cell, such as bacterial, insect, fungal, plant or animal cells. Preferred fungal cells are, for example, those of the genus Saccharomyces, in particular those of the species S. cerevisiae.

In addition, the present invention relates to a polypeptide encoded by the nucleic acid molecule obtained in accordance with the present invention. The proteins and polypeptides of the present invention are not necessarily translated from a designated nucleic acid sequence; the polypeptides may be generated in any manner, including for example, chemical synthesis, or expression of a recombinant

expression system, or isolation from a suitable viral system. The polypeptides may include one or more analogs of amino acids, phosphorylated amino acids or unnatural amino acids. Methods of inserting analogs of amino acids into a sequence are known in the art. The polypeptides may also include one or more labels, which are known to those skilled in the art. In this context, it is also understood that the proteins according to the invention may be further modified by conventional methods known in the art.

Diagnostic means and compositions

The invention also relates to a diagnostic composition comprising at least one of the aforementioned cell clone libraries and optionally suitable means for detection. Said diagnostic compositions can be preferably used for a method according to the present invention. The diagnostic composition of the invention may contain further ingredients such as selection markers and components for selective media suitable for the generation of transformed cells and transgenic non-human animals. Furthermore, the composition may include buffers and substrates for reporter genes that may be present in the vector of the invention employed in accordance with the method of the invention. The composition of the invention may advantageously be used for carrying out the method of the invention and could be, inter alia, employed in a variety of applications referred to herein, e.g., in the diagnostic field or as research tool. The parts of the diagnostic composition of the invention can be packaged individually in vials or in combination in containers or multicontainer units. Manufacture of the composition follows preferably standard procedures which are known to the person skilled in the art. The composition or its ingredients according to the invention can be used in mammalian cell and tissue cultures, for example, for any of the above described methods or for detecting inhibitors and activators of cell a given protein.

As is mentioned herein before, the method of the present invention is particular suited for breeding test animals, e.g., for pharmaceutical industry. The cell clones will be used to produce "knock-out" animals and hence avoid currently used method which are based on tedious homologous recombination techniques. Potentially

thousands of transgenic animals bearing deficience in a specific protein could be produced. A hallmark of the method will be the speed, high sensitivity and low cost.

Likewise, the present invention relates to the use of cell clone libraries obtained by the method of the present invention and the use of non-human animals described herein before for the screening of drugs. For example, cell clones or non-human animals that display a certain phenotype due to the expression of ribozyme gene can be screened for compounds that are capable of compensating said phenotype. The present invention thus relates also to a process for identifying therapeutically active compounds, said method comprising:

(a) contacting a cell clone or non-human animal of the invention with a compound to be screened; and

(b) determining if the compound compensates for the defect in the cell clone or animal.

Reversely, it can be determined whether compound has lost its biological effect because of the missing of its target in the cell clone.

Said compounds may be comprised in, for example, samples, cell extracts from, e.g. plants, animals or microorganisms. Furthermore, said compounds may be known in the art but hitherto not known to be an antagonist/inhibitor or agonist/activator of the gene(s) suppressed or activated by the ribozyme. Preferably said sample comprises substances of similar chemical and/or physical properties, most preferably said substances are identical. The compounds which can be prepared and identified according to the method of the present invention may be expression libraries, e.g., cDNA expression libraries, peptides, proteins, nucleic acids, antibodies, small organic compounds, ligands, hormones, peptidomimetics, PNAs or the like.

In a further embodiment, the present invention relates to the use of ribozymes for the method of the present invention as described herein before. As mentioned in the foregoing embodiments, any ribozyme may be used which can be directed to a given target nucleic acid molecule.

These and other embodiments are disclosed and encompassed by the description and examples of the present invention. Further literature concerning any one of the methods, uses and compounds to be employed in accordance with the present invention may be retrieved from public libraries, using for example electronic devices. For example the public database "Medline" may be utilized which is available on the Internet, for example under http://www.ncbi.nlm.nih.gov/PubMed/medline.html. Further databases and addresses, such as http://www.ncbi.nlm.nih.gov/, http://www.infobiogen.fr/, http://www.fmi.ch/biology/research_tools.html, http://www.tigr.org/, are known to the person skilled in the art and can also be obtained using, e.g., http://www.lycos.com. An overview of patent information in biotechnology and a survey of relevant sources of patent information useful for retrospective searching and for current awareness is given in Berks, TIBTECH 12 (1994), 352-364.

The present invention is further illustrated by reference to the following non-limiting examples.

Example

Unless stated otherwise in the examples, all recombinant DNA techniques are performed according to protocols as described in Sambrook et al. (1989), Molecular Cloning : A Laboratory Manual. Cold Spring Harbor Laboratory Press, NY or in Volumes 1, 2 of Ausubel et al. (1994), Current Protocols in Molecular Biology. Standard materials and methods for plant molecular work are described in Plant Molecular Biology Labfase (1993) by R.D.D. Croy, jointly published by BIOS Scientific Publications Ltd (UK) and Blackwell Scientific Publications (UK).

A combinatorial mixture of ribozymes was produced in a following way. Two oligonucleotides were synthesized 5'-GTCGGCGCGCCNNNNNNNNNNNNNCTGATGAGTCCGTGAGGACGAAA-3' (SEQ ID NO: 2) and 5'-GACTTAATTAANNNNNNNNNNNN*TTTCGTCCTCACGGACTCATCAG-3' (SEQ ID NO: 3). Where N is a mixture of A,C,G,T and N* is 10%A, 40%C, 40G% and 10%T. The oligos (25 µg of each) were incubated for 5 min at 85°C in 20mM Tris, pH 7.5

and 10mM $MgCl_2$ in final volume of 100 µl. After chilling to 65°C, 2mM dNTPs and 2.5 Units of Tac-polymerase were added. The incubation was performed for 30 min at 65°C. The reaction was followed by phenol/chloroform extraction. The obtained filled-in oligo was dissolved in 10 µl of ligation buffer (Boehringer Mannheim) and cut by restriction enzymes AscI and PacI (10 units of each) and subsequently extracted with phenol/chloroform. The resulted oligo was ligated with pGvaLber (see below) wich was cut with same restriction enzymes.

pGvaLber was constructed in a following way. pGvaL (see DE-C1 4424 761) was used as a template for nested PCR with following oligos:

1. 5'-GTCCTAGACCGTGCAAAAGGAGAGCC-3' (SEQ ID NO: 4)
2. 5'-GCACTAGTGCAGCAGCCGCCGCGCC-3' (SEQ ID NO: 5)
3. 5'-GGCGCGCCTCGAGTTAATTAATGCAGCCTGTGGACCCAACGAC-3' (SEQ ID NO: 6)
4. 5'-TTAATTAACTCGAGGCGCGCCGACGTCGCACACCTGGGTTCGA-3' (SEQ ID NO: 7)

Two PCR-reactions were done: first with oligos 1 and 4 (I) and the second with oligos 2 and 3 (II). The reaction volumes were 50 µl. Subsequently a third PCR-reaction was performed using as a template the reaction mixtures of I and II and as primer oligos 1 and 2. The overall PCR-product was ligated into a commercial vector pGEM-T(Promega). This construct was called pGvaLber.

As mentioned above pGvaLber was cut with AscI and PacI and ligated with the combinatorial mixture of ribozymes. The resulting library was called pGvalRzber.

After transfection of competent bacteria (TOP10, Invitrogen) with electroporation about 10^9 independent clones per µg of DNA were obtained.

For preparation of ribozyme libraries in other vectors pOPRSVI-1 or pTV-0 the overall PCR-product was cut with restriction enzymes XbaI ans SpeI and ligated into corresponding plasmids using the XbaI sites. The correct orientation of the insert was tested by sequencing. The resulting vectors are called correspondingly pOPRSVI-1-VaLRz and pTV-0-VaLRz.

All general cloning procedures were performed according to: Sambrook, et al. (1989). Molecular cloning. Cold Spring Harbor laboratory, Cold Spring Harbor, NY.

Example 1:

Madin-Darby Canine Kidney (MDCK, ATCC-collection #CCL-34) cells were cultivated in DMEM and 10% calf serum. About 10^7 MDCK-cells on eight 15cm Petri dishes were transfected with 100 µg pOPRSVI-1-VaLRz using a standart Ca-precipitation protocol, Current Protocols in Cell Biology, John Wiley and Sons, New York (1997). The permanently transfected cells were selected in a medium containing 500 µg/ml of G418 for ten to fourteen days. Cells were trypsinized, harvested, resuspended in total of 1ml of PBS and injected into 6 nude mice. After 18 days animals were sacrificed and analysed for the formation of visible tumours. Three of the mice developed tumours of about 1-2cm. Non-transfected MDCK cells developed no tumours upon injection into control animals. Total RNA prepared from these tumours was used to produce cDNA (Sambrook et al, see supra). Specific ribozymes were amplified by PCR-technique using as primers oligos flanking ribozyme in the vector: The amplified fragment of about 100bp was ligated into pGEM-T vector. 50 independent clones were sequenced. Amplification of cDNA from one tumour was not successful, two other tumours gave in sum 14 different sequences two of which are part of known suppressor gene.

Example 2:

A combinatorial ribozyme library inserted in pTV-0 vector (pTV-0-VaLRz) was used to transfect 10^6 mouse embryonic stem cell (ES, ATCC collection, #CRL-2239). After transfection clones were selected for neo-resistance. The transfection and the selection procedures were the same as described in Example 1. About 60 clones were obtained after 12 days. Total RNA prepared from each of these clones was used to produce cDNA. Specific ribozymes were amplified by PCR-technique using as primers oligos flanking ribozyme in the vector (see above Example 1). The amplified fragment of about 100bp was ligated into pGEM-T vector and subsequently sequenced.

Example 3:

CHO-cells (ATCC-collection, # CCL-61) were transfected with pOPRSVI-1-VaLRz and permanent transfected cells were selected as described above for MDCK cells. About 10^6 cells were incubated with a mixture of BODIPY-Ceramide (Molecular Probes, USA) and defatted BSA in serum free DMEM, 5 µM of each, for 1 hour at 37°C. After trypsinization cells were subjected to cell sorting in Becton-Dickinson machine and 5% of most fluorescent cells were collected. Cells were allow to grow for another 3 days and the incubation with BODIPY-ceramide and subsequent cell sorting was repeated. Obtained cells were seeded sparsely to produce colonies. Preparation of cDNA, amplification of ribozymes and sequencing was performed as described above. This application should give mutant with defect intracellular lipid/protein transport.

CLAIMS

1. A method for producing cell clone libraries comprising the steps of:
   (a) transfecting mammalian cells with a pool of expression vectors comprising
      (i) ribozyme genes which comprise a central catalytic domain and flanking sequences of at least 6 nucleotides in length each of random base sequence and, optionally,
      (ii) a selection marker;
   (b) cultivating the transfected cells under conditions which permit expression of the ribozyme gene present in the vectors;
   (c) identifying those cell clones expressing the ribozyme;
   (d) optionally subdividing the vector pool(s) identified in step (c) and repeating step (a) to (c); and
   (e) examining the so-identified cell clones for a modification of the properties they originally had.

2. The method of claim 1, wherein the mammalian cells are MDCK cells, CHO cells, or embryonic stem cells.

3. The method of claim 1 or 2, wherein the regulatory elements controlling the expression of the ribozyme genes comprise a constitutive, tissue-specific, organ-specific or inducible promoter.

4. The method of any one of claims 1 to 3, wherein the selection marker is an antibiotics selection marker or a reporter gene.

5. The method of any one of claims 1 to 4, wherein the reporter gene encodes a luciferase.

6. The method of any one of claims 1 to 4, wherein the antibiotics selection marker encodes neomycin phosphotransferase.

7. The method of any one of claims 1 to 6, wherein the selected properties are detected by fluorescence analysis and the cells are optionally enriched by cell sorting.

8. The method of any one of claims 1 to 7, wherein the flanking sequences in the ribozyme gene are 6 to 13 nucleotides in length.

9. The method of any one of claims 1 to 8, wherein the catalytic domain in the ribozyme gene consists of a hammerhead structure.

10. The method of claim 9, wherein said hammerhead structure has the sequence 5'-CTGATGAGTCCGTGAGGACGAAAAC-3'.

11. The method of any one of claims 1 to 10, wherein the ribozyme is permanently expressed in cell clones.

12. The method of any one of claims 1 to 11, wherein the phenotype-determining ribozyme in the selected cell clone is amplified by (RT-)PCR.

13. The method of claim 12, wherein said ribozyme is identified by sequence analysis.

14. The method of claim 12 or 13, wherein said ribozyme is cloned into a vector.

15. The method of any one of claims 12 to 14, wherein the target nucleic acid molecule(s) of the phenotype-determining ribozyme is deduced from the flanking sequences of said ribozyme.

16. The method of claim 15, wherein said nucleic acid molecule(s) are identified by data bank analysis or by cloning from DNA libraries.

17. The method of claim 16, wherein said DNA library is a cDNA library.

18. The method of any one of claims 15 to 17, wherein said target nucleic acid molecule encodes a protein.

19. The method of claim 18, wherein said protein is a tumor suppressor or an intracellular lipid-protein transporter protein.

20. The method of any one of claims 1 to 19, wherein said cultivation in step (b) is performed in an animal.

21. The method of claim 20, wherein said animal is mice.

22. Cell clone libraries obtainable by the method of any one of claims 1 to 21.

23. The cell clone library of claim 22, wherein said cells consist of embryonic stem cells.

24. Non-human animal comprising the embryonic stem cells of claim 23.

25. The non-human animal of claim 24 which is mice.

26. A nucleic acid molecule obtainable by a method of any one of claims 15 to 21 which encodes a protein having enzymatic, binding and/or regulatory activity.

27. A nucleic acid molecule which specifically hybridizes with the nucleic acid molecule of claim 26 and which encodes a mutated version of said protein which has lost its enzymatic, binding and/or regulatory activity.

28. A vector comprising the nucleic acid molecule of claim 26 or 27.

29. A host cell comprising the vector of claim 28.

30. A polypeptide encoded by the nucleic acid molecule of claim 26 or 27.

31. A diagnostic composition comprising the cell clone library of claim 22 or 23 and optionally suitable means for detection.

32. Use of a method of any one of claims 1 to 21, for breeding test animals.

33. Use of the cell clone library of claim 22 or 23 or a non-human animal of claim 24 or 25 for drug screening.

34. Use of ribozymes for a method of any one of claims 1 to 21.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V.
(C) CITY: Berlin
(E) COUNTRY: DE

(ii) TITLE OF INVENTION: Process for producing cell clone libraries

(iii) NUMBER OF SEQUENCES: 7

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CTGATGAGTC CGTGAGGACG AAAAC 25

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GTCGGCGCGC CNNNNNNNNN NNNNCTGATG AGTCCGTGAG GACGAAA 47

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 46 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GACTTAATTA ANNNNNNNNN NNNTTTCGTC CTCACGGACT CATCAG 46

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GTCCTAGACC GTGCAAAAGG AGAGCC 26

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GCACTAGTGC AGCAGCCGCC GCGCC 25

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GGCGCGCCTC GAGTTAATTA ATGCAGCCTG TGGACCCAAC GAC 43

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

TTAATTAACT CGAGGCGCGC CGACGTCGCA CACCTGGGTT CGA 43

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 5/10, 9/00, 15/10, A01K 67/027

A3

(11) International Publication Number: **WO 99/53031**

(43) International Publication Date: 21 October 1999 (21.10.99)

(21) International Application Number: PCT/EP99/02391

(22) International Filing Date: 8 April 1999 (08.04.99)

(30) Priority Data:
198 15 590.5 8 April 1998 (08.04.98) DE

(71) Applicant *(for all designated States except US)*: MAX-PLANCK-GESELLSCHAFT ZUR FÖRDERUNG DER WISSENSCHAFTEN E.V. [DE/DE]; Berlin (DE).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: KURZCHALIA, Teymuras [GE/DE]; Fehrbelliner Strasse 97, D-10119 Berlin (DE).

(74) Agent: VOSSIUS & PARTNER; Siebertstrasse 4, D-81675 München (DE).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
16 December 1999 (16.12.99)

(54) Title: PROCESS FOR PRODUCING CELL CLONE LIBRARIES

(57) Abstract

Described is a process for producing cell clone libraries via the use of a pool of expression vectors comprising ribozyme genes.Furthermore, cell clones with modified phenotypes and non-human animals derived from such cell clones are provided as well as novel genes identified from the cell clone library. Fields of the application of the invention are molecular-biological research and the pharmaceutical industry.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

INTERNATIONAL SEARCH REPORT

Intr *tional Application No
PCT/EP 99/02391

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N5/10 C12N9/00 C12N15/10 A01K67/027

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DE 44 24 762 C (MAX PLANCK GESELLSCHAFT) 27 July 1995 (1995-07-27) abstract column 2, line 45 -column 6, line 10; claim 1 ---	1,2, 8-10, 12-18, 22,26, 28-30
X	LARSSON S ET AL., : "Reduced beta2-microglobulin mRNA levels in transgenic mice expressing a designed hammerhead ribozyme" NUCLEIC ACIDS RESEARCH, vol. 22, no. 12, 1994, pages 2242-2248, XP002120162 the whole document --- -/--	1-3,8,9, 11,12, 18, 22-26, 28-30,34

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 October 1999	05/11/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Mateo Rosell, A.M.

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Int'l Application No
PCT/EP 99/02391

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 93 11230 A (DYNAL AS) 10 June 1993 (1993-06-10) page 1, paragraph 1; example 2	1,2,4,8, 18,26, 28-30
A	WO 94 13833 A (INNOVIR LAB INC) 23 June 1994 (1994-06-23) page 5, paragraph 4; examples 1,2 page 12, paragraph 1	1,9,12, 14,26, 28-30
A	WO 97 27213 A (UNIV LELAND STANFORD JUNIOR) 31 July 1997 (1997-07-31) page 3, line 1-27 page 4, line 26 -page 5, line 10 page 28, line 19 -page 34, line 8 page 37, line 25 -page 38, line 16	1,7-9, 18,19, 26, 28-30,33
A	SKRYABIN BV AND SCHMAUSS C: "Enhanced selection fro homologous-recombinant embryonic stem cell clones with a neomycin phosphotransferase gene in antisense orientation" TRANSGENIC RESEARCH, vol. 6, 1997, pages 27-35, XP002120163 the whole document	1-4,6,8, 9,11,18, 22,23, 25,28-30
A	XIE Y F ET AL: "A ribozyme -mediated, gene ''knockdown'' strategy for the identification of gene function in zebrafish" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 94, no. 25, 9 December 1997 (1997-12-09), pages 13777-13781, XP002108912 ISSN: 0027-8424	1,2,8,9, 11,14, 20, 27-30,34
A	JARVIS T C ET AL: "INHIBITION OF VASCULAR SMOOTH MUSCLE CELL PROLIFERATION BY HAMMERHEAD RIBOZYMES TARGETTING C-MYB" JOURNAL OF CELLULAR BIOCHEMISTRY, vol. 19A, 1 January 1995 (1995-01-01), page 221 XP002024063 ISSN: 0730-2312 abstract	1,3,9, 18, 26-30, 33,34

-/--

1

Int ational Application No
PCT/EP 99/02391

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	HASELOFF J ET AL: "SIMPLE RNA ENZYMES WITH NEW AND HIGHLY SPECIFIC ENDORIBONUCLEASE ACTIVITIES" NATURE, vol. 334, 18 August 1988 (1988-08-18), pages 585-591, XP002029776 ISSN: 0028-0836 cited in the application the whole document	1
P,X	WO 98 32880 A (BARBER JACK ;IMMUSOL INC (US); TRITZ RICHARD (US); WELCH PETER (US)) 30 July 1998 (1998-07-30) abstract page 4, paragraph 1 -page 8, last paragraph page 65, paragraph 2 -page 81, last paragraph	1-4,6,7, 11,12, 18,19, 26, 28-30,34

INTERNATIONAL SEARCH REPORT

information on patent family members

Int[illegible]tional Application No
PCT/EP 99/02391

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
DE 4424762 C	27-07-1995	WO 9601314 A	18-01-1996
		EP 0776363 A	04-06-1997
WO 9311230 A	10-06-1993	AU 3084492 A	28-06-1993
WO 9413833 A	23-06-1994	AU 675482 B	06-02-1997
		AU 5739694 A	04-07-1994
		EP 0681613 A	15-11-1995
		JP 8507202 T	06-08-1996
		US 5589332 A	31-12-1996
WO 9727213 A	31-07-1997	AU 1707497 A	20-08-1997
		AU 1707897 A	20-08-1997
		CA 2244222 A	31-07-1997
		EP 0877752 A	18-11-1998
		WO 9727212 A	31-07-1997
WO 9832880 A	30-07-1998	AU 5927598 A	18-08-1998